

# Quantum Cryptography with Local Bell Tests

Charles Ci Wen Lim,<sup>1</sup> Christopher Portmann,<sup>1,2</sup> Marco Tomamichel,<sup>2,3</sup> Renato Renner,<sup>2</sup> and Nicolas Gisin<sup>1</sup>

<sup>1</sup>*Group of Applied Physics, University of Geneva, Switzerland.*

<sup>2</sup>*Institute for Theoretical Physics, ETH Zurich, Switzerland.*

<sup>3</sup>*Centre for Quantum Technologies, National University of Singapore, Singapore.*

One must be careful when assessing the security of practical quantum key distribution systems as real devices do not necessarily comply with the assumed theoretical models. In fact, a complete characterization of the devices is required by most security proofs—which is a non-trivial task in practice. In this work, we propose a quantum cryptography protocol that only requires partial knowledge of the devices. Specifically, the security of the protocol is based on a novel self-testing procedure where Alice and Bob perform Bell tests independently of each other. Then, by establishing a relation between the local Bell tests and a recent entropic uncertainty relation, we show that the protocol is secure against general attacks in the finite key size regime.

**Introduction.** Quantum key distribution (QKD) [1] is a cryptographic technology which allows two legitimate users (traditionally called Alice and Bob) to generate provably secure cryptographic keys. Its prowess derives from the fact that Alice and Bob can perform tests to detect possible attacks; a feat exclusive to quantum cryptography. However, the requirements for such tests are rather demanding, i.e., the involved devices must conform to certain theoretical models, which in practice is generally hard to achieve. In fact, failure to do so may open certain side-channels [2], which an adversary can exploit without the risk of being detected. To put it differently, it is essential to have complete knowledge of the devices, otherwise most security proofs do not apply. On the other hand, complete characterization of devices usually necessitates additional parameters—to characterize the discrepancies between the devices and the theoretical models—and as a result, the secret key rate of a practical QKD scheme may depend on several parameters, e.g., see Ref [3]. Clearly, this scenario is not encouraging for QKD systems with finite resources<sup>1</sup>: if the number of discrepancies is large, then the additional parameters required to characterize the discrepancies, together with its statistical fluctuations, are likely to penalize the secret key rate [4–10].

In light of this dilemma, it is useful to consider the converse problem: instead of designing devices that conform to the theoretical models, we want to devise security proofs that are valid for a very general class of devices which nonetheless can be characterized by very few parameters. For instance, such a security proof was made available by Ref [10] where the devices only need to be characterized by the *overlap* of the measurements [11–13]. However, the knowledge of the overlap implies that Alice and Bob have complete knowledge of the measurement devices or they make the assumption that the measurement devices conform to the theoretical models.

In this work, we propose a QKD protocol that is based on a novel local self-testing procedure [14]. In particular,

the devices are tested locally, i.e, Alice and Bob perform the Clauser-Horne-Shimony-Holt (CHSH) test [15], an application of Bell’s theorem [16] on their own devices, independent of each other and the quantum channel. As a result, a complete characterization of the devices is not necessary for the security of the protocol. Furthermore, the CHSH test is independent of the quantum channel and thus the channel loss cannot be used to open the detection loophole [17]. Contrary to most QKD protocols, the protocol adopts the tripartite model of Ref [18–21] where Alice and Bob generate and send quantum states to Charlie whose task is to perform an entangling measurement (similar to entanglement swapping [22]) on it. Then, the security assessment of the quantum channel (including Charlie) follows the channel estimation technique of the BB84 protocol [23], i.e., checking for errors in the bases X and Z. In addition, we make the following assumptions<sup>2</sup> on the laboratories of Alice and Bob: A1) access to trusted local sources of randomness, A2) access to an authenticated, but otherwise insecure classical channel, A3) no information is allowed to leave the laboratories unless the protocol prescribes it, A4) access to trusted classical operations, A5) the measurement and source devices do not have internal memories and A6) the marginal states of Alice and Bob are independent of whether Charlie outputs a pass or fail.

Under the above assumptions, and by deriving a relation between the CHSH test and a recent security proof technique (based on an entropic uncertainty relation for smooth entropies [13]), the security proof in the finite key size regime is obtained. Moreover, our result is intuitively related to the almost tight finite-key analysis [10] of the BB84 protocol and it differs only by a term that is dependent on the CHSH value. Most importantly, using realistic CHSH values, we obtained secret key rates that are comparable to the ones of the BB84 protocol.

**Related Work.** Device-independent QKD [24–28] whose security is based on the monogamy of non-local

<sup>1</sup> Alice and Bob exchange finite number of systems in the protocol.

<sup>2</sup> A full discussion on the assumptions is detailed in the Supplementary Material.

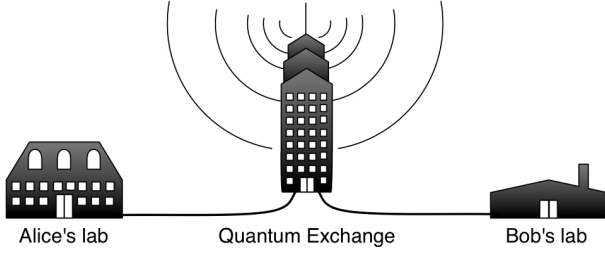


FIG. 1. **Quantum exchange.** Alice and Bob each send a quantum state to Charlie. Then, Charlie is supposed to make an entangling measurement on the quantum states, and if it is successful he outputs a pass, otherwise he outputs a fail. In addition, whenever the entangling measurement is successful, Charlie broadcasts another two more bits such that either Alice or Bob can make the correction bit-flip operation if required.

correlations [29] offers the same advantage, i.e., there is no need to have complete knowledge of the devices. The main difference between our protocol and device-independent QKD lies in the security principles which is reflected in the application of the CHSH test, i.e., we use it to test the devices locally while device-independent QKD use it to test both the quantum channel and the devices. Although our protocol is more intricate than the original version of device-independent QKD, it actually offers two advantages. First, the local CHSH test can be made loophole-free more readily than the CHSH test of device-independent QKD. Second, in the limiting case where the CHSH value is maximal, the secret key rate of our protocol is completely independent of the channel loss.

**Basic ideas.** We start by defining the network topology of the protocol (motivated by Ref [18–21]): Alice and Bob prepare certain quantum states and send them to a quantum exchange (see Figure 1)—akin to a telephone exchange that connects phone calls. Here, the task of the quantum exchange is to establish entanglement between Alice and Bob such that they can use it for quantum cryptography. In the ideal case, Alice and Bob each randomly prepare one of the BB84 states and send it to Charlie who makes a Bell state measurement; which is equivalent to Charlie distributing Bell states and Alice and Bob measuring it in the computational and diagonal bases [18, 19]. Note that Charlie needs to inform Alice and Bob which of the four Bell states he obtained, so that one of them flips a bit of their outcome if required. Then, it can be easily verified that after the correction bit-flip operations, the bit strings of Alice and Bob are perfectly correlated. In the event that the Bell state measurement is unsuccessful, he outputs a fail, e.g., because of losses (for more details, we refer to Ref [18, 19]). Interestingly, the idea of using entanglement swapping for quantum cryptography can be utilized to rule out all Trojan-horse attacks [30] on the laboratories of Alice and Bob, which is a highly desirable feature for practical quantum cryp-

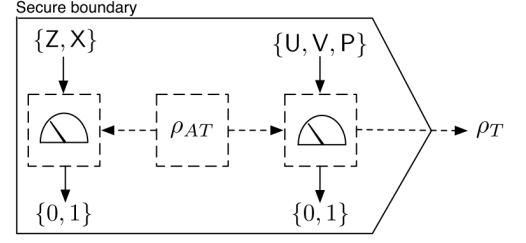


FIG. 2. **Arrangement of devices.** Alice/Bob has three devices i.e., a source device and two measurement devices, of which they have minimal knowledge. By arranging the devices according to the self-testing procedure, Alice/Bob can either perform the CHSH test or generate an output state  $\rho_T$ .

tography. Moreover, the security holds even if Charlie is untrusted. In fact, as recently shown by Ref [20, 21], such protocols can be extended to more general settings.

Although the above protocols [18–21] have an important advantage—avoiding Trojan-horse attacks—over traditional protocols, they still require accurate description of the source devices. For instance, the protocols of Ref [18, 19] require that the source devices produce precisely the BB84 states. However, as we mentioned earlier, such a requirement may entail undesirable demands on the resources. In the following, we adopt the self-testing procedure of Ref [14] to overcome it, i.e., the functionality of the source devices do not need to adhere to such strict requirements.

We briefly discuss the self-testing procedure for Alice, but the same holds for Bob. Alice holds three devices: a source device that claims to produce bipartite maximally entangled states and two measurement devices. The first measurement device has two settings  $\{Z, X\}$ <sup>3</sup> with binary outputs and the second measurement has three settings  $\{U, V, P\}$  where the first two settings produce binary outputs and the last setting sends the other half of the bipartite state to Charlie. By arranging the devices according to the configuration of Ref [14] (see Figure 2), Alice has two choices, namely she can either select  $P$  and let one half of the bipartite state be sent to Charlie or use the settings  $U, V$  to perform the CHSH test. We refer to the former as sub-protocol  $\Gamma_{\text{QKD}}$  and the latter as sub-protocol  $\Gamma_{\text{CHSH}}$ , and the formal descriptions are given in the protocol definition section.

The basic idea of the CHSH test is made clear with the following observations. First, according to the entropic uncertainty relations [11–13], if the measurements corresponding to bases  $X$  and  $Z$  do not commute, then secret key distillation is possible. Second, from Ref [31] we know that only non-commuting measurements can be used to violate the CHSH test. In fact, the maximum violation of the CHSH test requires anti-commuting

<sup>3</sup> We label the measurements  $Z$  and  $X$  because measurements in the computational and diagonal basis are optimal, but these can be arbitrary.

measurements. Putting these observations together, it is easy to see that the CHSH test is a device-independent test of the uncertainty principle and thus can be used to prove the security of the protocol. More precisely, we bound the overlap of the measurements corresponding to bases  $X$  and  $Z$  with the observed CHSH value, which in turn allows us to apply a generalized version of the entropic uncertainty relation for smooth entropies [13] to the protocol.

*Note:* Such a relation has also been obtained independently by Ref [14] but with a different proof method.

**Security definition.** It is instructive to briefly recall the criteria for a QKD protocol to be secure

First, we fix some notations. Let  $S_A$  and  $S_B$  be the key strings of Alice and Bob, respectively, and let  $E$  be the information which the adversary gathers over the execution of the protocol. Then, the joint state of  $S_A$  and  $E$  can be described as a classical-quantum state,  $\rho_{S_A E} = \sum_s |s\rangle\langle s| \otimes \rho_E^s$  where  $\{\rho_E^s\}_s$  are states held by the adversary. We recall that a secure QKD protocol meets two requirements [10], namely correctness and secrecy. Correctness is achieved if  $S_A = S_B$  and secrecy is achieved if  $\rho_{S_A E} = U_{S_A} \otimes \rho_E$  where  $U_{S_A}$  is the uniform mixture of all possible values of the key string. The former means that the key strings of Alice and Bob are identical and the latter implies that the key string of Alice is uniformly distributed and independent of  $E$ .

However, in practice, even in the best scenario, a minuscule amount of errors are inevitable. Thus, for small parameters  $\varepsilon_{\text{cor}}$  and  $\varepsilon_{\text{sec}}$ , we require that the protocol is  $\varepsilon_{\text{cor}}$ -correct, i.e.,  $\Pr[S_A \neq S_B] \leq \varepsilon_{\text{cor}}$ , and  $\varepsilon_{\text{sec}}$ -secret, i.e.,  $\frac{1}{2} \|\rho_{S_A E} - U_{S_A} \otimes \rho_E\|_1 \leq \varepsilon_{\text{sec}}$ . Furthermore, such a security definition guarantees that the protocol is universally composable [32], i.e., the secrecy of  $S_A$  w.r.t  $E$  holds even when  $S_A$  is used in other cryptographic schemes.

**Protocol definition.** The protocol is characterized by a set of field specifications  $\Phi(\ell, m_x, m_z, j, S_{\text{tol}}, Q_{\text{tol}}, \text{leak}_{\text{EC}}, \varepsilon_{\text{cor}})$  which is parameterized by the secret key length  $\ell$ , the classical post-processing block size  $m_x$ , error rate estimation sample size  $m_z$ , CHSH test sample size  $j$ , tolerated CHSH value  $S_{\text{tol}}$ , tolerated channel error rate  $Q_{\text{tol}}$ , error correction leakage  $\text{leak}_{\text{EC}}$  and the required correctness  $\varepsilon_{\text{cor}}$ .

*1. State preparation and distribution.*— Alice selects a sub-protocol  $h_i \in \{\Gamma_{\text{QKD}}, \Gamma_{\text{CHSH}}\}$  where  $\Gamma_{\text{QKD}}$  is selected with probability  $1-p_s$  and  $\Gamma_{\text{CHSH}}$  with probability  $p_s$ . In the following, we describe sub-protocols  $\Gamma_{\text{QKD}}$  and  $\Gamma_{\text{CHSH}}$  formally for each  $i$ th run

- $\Gamma_{\text{QKD}}$ : Alice selects a measurement setting  $a_i \in \{X, Z\}$  with probabilities  $p_x$  and  $1-p_x$ , respectively, measures one half of the bipartite state with it and stores the measurement output in  $y_i$ . The other half of the bipartite system is sent to Charlie.

- $\Gamma_{\text{CHSH}}$ : Alice measures both halves of the bipartite state: she chooses two bit values  $u_i, v_i$  uniformly at random, where  $u_i$  sets the measurement on the first half to  $X$  or  $Z$  and  $v_i$  sets the measurement on the second half to  $U$  or  $V$ . The outputs of each measurement are recorded in  $s_i$  and  $t_i$ , respectively.

Similarly, Bob records his choice of sub-protocol in  $h'_i$  and his measurement settings and outputs for sub-protocols  $\Gamma_{\text{QKD}}$  and  $\Gamma_{\text{CHSH}}$  in  $b_i, y'_i$  and  $u'_i, v'_i, s'_i, t'_i$ , respectively.

*2. Quantum exchange.*— Charlie makes an entangling measurement on the quantum states sent by Alice and Bob, and if it is successful, he broadcasts  $f_i = \text{pass}$ , otherwise he broadcasts  $f_i = \text{fail}$ . Furthermore, if  $f_i = \text{pass}$ , then Charlie communicates  $g_i \in \{0, 1\}^2$  to Alice and Bob. Then either Alice or Bob flips a bit of their corresponding measurement outcome if required.

*3. Sifting.*— Alice and Bob announce their sub-protocol and basis choices  $\{h_i\}_i, \{h'_i\}_i, \{a_i\}_i, \{b_i\}_i$  over an authenticated classical channel and identify four sets,

- Key generation,  $\mathcal{X} := \{i : (h_i = h'_i = \Gamma_{\text{QKD}}) \wedge (a_i = b_i = X) \wedge (f_i = \text{pass})\}$
- Channel error rate estimation,  $\mathcal{Z} := \{i : (h_i = h'_i = \Gamma_{\text{QKD}}) \wedge (a_i = b_i = Z) \wedge (f_i = \text{pass})\}$
- Alice and Bob CHSH test sets,  $\mathcal{J} := \{i : h_i = \Gamma_{\text{CHSH}}\}$  and  $\mathcal{J}' := \{i : h'_i = \Gamma_{\text{CHSH}}\}$ , respectively.

The protocol repeats steps (1)-(3) as long as  $|\mathcal{X}| < m_x$  or  $|\mathcal{Z}| < m_z$  or  $|\mathcal{J}| < j$  or  $|\mathcal{J}'| < j$ , where  $m_x, m_z, j \in \mathbb{N}_1$ . We refer to these conditions as the sifting condition.

*4. Parameter estimation.*— To compute the average CHSH value from  $\mathcal{J}$ , Alice uses the following formula,  $S_{\text{test}} := 8 \sum_{i \in \mathcal{J}} f(u_i, v_i | s_i, t_i) / |\mathcal{J}| - 4$ , where  $f(u_i, v_i | s_i, t_i) = 1$  if  $s_i \oplus t_i = u_i \wedge v_i$ , otherwise  $f(u_i, v_i | s_i, t_i) = 0$ . Similarly, Bob uses the same formula and arrives at  $S'_{\text{test}}$ . Next, both Alice and Bob publicly announce the corresponding bit strings  $\{y_i\}_{i \in \mathcal{Z}}, \{y'_i\}_{i \in \mathcal{Z}}$  and compute the average error rate  $Q_{\text{test}} := \sum_{i \in \mathcal{Z}} y_i \oplus y'_i / |\mathcal{Z}|$ . If  $\max\{S_{\text{test}}, S'_{\text{test}}\} < S_{\text{tol}}$  or  $Q_{\text{test}} < Q_{\text{tol}}$ , they abort the protocol.

*5. One-way classical post-processing.*— Alice and Bob choose a random subset of size  $m_x$  of  $\mathcal{X}$  for classical post-processing, and we let  $X$  and  $X'$  be random variables that take the values from the corresponding strings  $\{y_i\}_i$  and  $\{y'_i\}_i$ . Then, an information reconciliation scheme is applied, revealing at most  $(\text{leak}_{\text{EC}} + \lceil \log(1/\varepsilon_{\text{cor}}) \rceil)$ -bits of information. More specifically, an error correction scheme which leaks at most  $\text{leak}_{\text{EC}}$ -bits of information is applied, then an error verification scheme which leaks  $\lceil \log(1/\varepsilon_{\text{cor}}) \rceil$ -bits of information is applied. If the error verification fails, they abort the protocol. Finally, Alice

and Bob apply privacy amplification to their bit strings to obtain a secret key of length  $\ell$ .

**Security analysis and discussions.** In this section, we state our main result and discuss its feasibility.

$$\ell \leq m_x \left( 1 - \log \left( 1 + \frac{\hat{S}_{\text{tol}}}{4} \sqrt{8 - \hat{S}_{\text{tol}}^2} + \zeta \right) - h(\hat{Q}_{\text{tol}}) \right) - \text{leak}_{\text{EC}} - \log \frac{2}{\varepsilon_{\text{cor}}} - \log \frac{2}{\varepsilon^4}, \quad (1)$$

where  $h$  denotes the binary entropy function,  $\log$  denotes logarithm base 2,  $\hat{S}_{\text{tol}} := S_{\text{tol}} - \xi$  and  $\hat{Q}_{\text{tol}} := Q_{\text{tol}} + \mu$  with the statistical fluctuations given by

$$\xi := \sqrt{\frac{32}{j} \ln \frac{1}{\varepsilon}}, \quad \zeta := \sqrt{\frac{2(m_x + j)(j + 1)}{m_x j^2} \ln \frac{1}{\varepsilon}}, \quad \mu := \sqrt{\frac{(m_x + m_z)(m_z + 1)}{m_x m_z^2} \ln \frac{1}{\varepsilon}}. \quad (2)$$

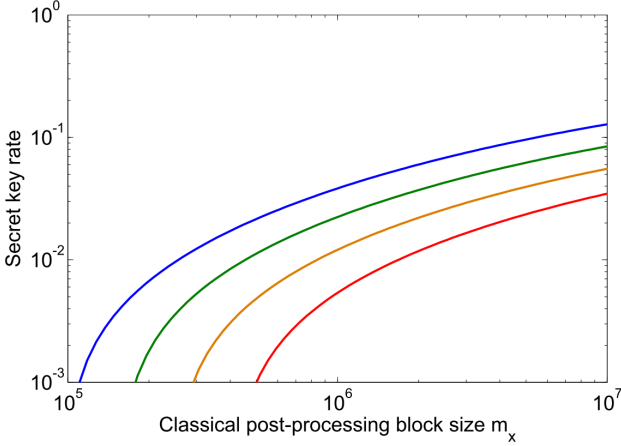


FIG. 3. **Secret key rate as a function of classical post-processing block size.** For a fixed channel error rate of  $Q_{\text{tol}} = 1.5\%$ , we plot the secret key rate for  $S_{\text{tol}} \in \{2.825, 2.800, 2.775, 2.750\}$  from left to right. The security bound  $\varepsilon_{\text{sec}}$  and required correctness are fixed to  $10^{-9}$  and  $10^{-12}$ , respectively.

The secret key rate is defined as  $R(\Phi) := \ell/N$  where  $N$  is the number of bipartite states which Alice generates for the protocol. The proof for a slightly more general result is provided in the Supplementary Material.

Prior to the discussion, it is important to briefly mention the problem of local losses (low efficient detectors, coupling losses, etc), namely the detection loophole. To simplify the discussion, we consider the case whereby low efficient detectors are used and the rest of the local losses are negligible. Then, to overcome the detection loophole, we can either trust the efficiency of the detectors or demand that the overall detection efficiency is sufficiently

high. The correctness of the protocol is determined by the error verification scheme which is parameterized by the required correctness  $\varepsilon_{\text{cor}}$ . Then for field specifications  $\Phi(\ell, m_x, m_z, j, j', S_{\text{tol}}, Q_{\text{tol}}, \text{leak}_{\text{EC}}, \varepsilon_{\text{cor}})$ ,  $\varepsilon > 0$  and  $\varepsilon_{\text{sec}} = 6\varepsilon$ , the protocol is  $\varepsilon_{\text{sec}}$ -secret if

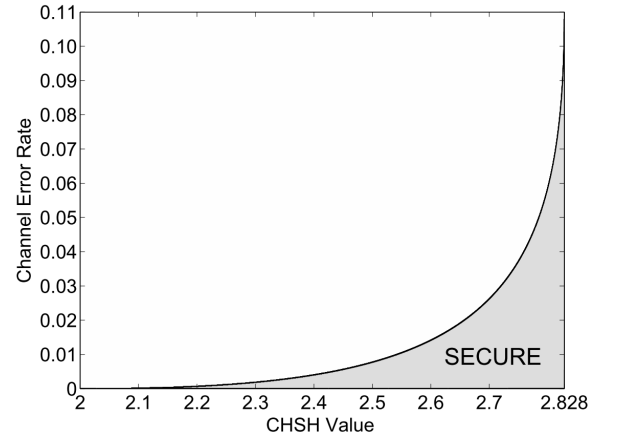


FIG. 4. **The secure region in the asymptotic limit as a function of CHSH value and channel error rate.** The maximum critical channel error rate is 11% which is the same as the BB84 protocol.

high<sup>4</sup>. For simplicity, we take Alice's and Bob's detectors to have perfect efficiency for our calculations, i.e., the measurement outcomes are restricted to binary outcomes.

First, we investigate the performance of the protocol with respect to the classical post-processing block size  $m_x$ . To proceed, the quantum channel is represented by a depolarizing channel parameterized by  $Q_{\text{tol}}$ . Next, we fix the security bound  $\varepsilon$  and optimize the secret key rate  $R(\Phi)$  over all field parameters  $\Phi$  that is  $\varepsilon$ -secure and has block size  $m_x$ . For the optimization, we set  $\text{leak}_{\text{EC}} := m_x f_{\text{EC}} h(\hat{Q}_{\text{tol}})$  with  $f_{\text{EC}} = 1.1$  where  $f_{\text{EC}}$  is the error correction efficiency (in the Shannon limit, it

<sup>4</sup> Note that detectors with an efficiency of 95% with negligible noise [33] were already reported

approaches 1). In Figure 3, we observe that reasonably good secret key rates can be achieved in the regime of  $m_x = 10^5$  bits, a factor of 10 larger than the finite key results [10] of BB84 protocol. Note that there were already QKD field tests [34] which worked on classical post-processing block size in the order of  $10^6$  bits and demonstrations [35] of CHSH tests that achieve violations from about 2.73 to 2.81.

In the asymptotic limit, that is,  $N \rightarrow \infty$ , and given there is no channel loss, i.e., Charlie always outputs a pass, it is easy to verify that  $m_x/N \rightarrow 1$  and the secret key rate reaches

$$R^\infty(\Phi) = 1 - \log \left( 1 + \frac{S_{\text{tol}}}{4} \sqrt{8 - S_{\text{tol}}^2} \right) - 2h(Q_{\text{tol}}). \quad (3)$$

Here, one can immediately see the roles of the sub-protocols  $\Gamma_{\text{CHSH}}$  and  $\Gamma_{\text{QKD}}$ : the local CHSH tests estimate the quality of the devices and the bit error rate estimates the quality of the quantum channel. In the case of  $S_{\text{tol}} = S_{\text{test}} = 2\sqrt{2}$ , we recover the asymptotic secret key rate [36] of the BB84 protocol.

**Conclusion.** In this work, we propose a QKD protocol based on local CHSH tests where the precise specification of the devices is not necessary. Then, by deriving a relation between the CHSH test and a generalized version of the entropic uncertainty relation, a security proof which is valid in the finite-key region is obtained. Most importantly, with realistic field parameters, the secret key rates are comparable to the ones of the BB84 protocol. Furthermore, the local CHSH tests can be readily made loophole-free as compared to device-independent QKD.

**Acknowledgments.** We thank Jean-Daniel Bancal, Nicolas Sangouard, Marcos Curty, Stefano Pironio, Hugo Zbinden and Valerio Scarani for helpful discussions. We acknowledge support from the National Centre of Competence in Research QSIT, the Swiss NanoTera project QCRYPT, the Swiss National Science Foundation (grant no. 200020-135048), FP7 Marie-Curie IAAP QCERT project and the European Research Council (grant No. 258932 and No. 227656).

---

## SUPPLEMENTARY MATERIAL

In the following, we provide the security proof for the protocol described in the main text. First, we define the assumptions required for the proof and then introduce the necessary technical lemmas. Second, we prove a relation between the CHSH test and a generalized version of smooth entropic uncertainty relation (Lemma 5). Third, we provide the required statistical statements for estimating certain quantities of the raw key. Finally, we state our main result (Theorem 1) which is slightly more general than the result given in (1).

### 1. Notations

We assume that all Hilbert spaces denoted by  $\mathcal{H}$ , are finite-dimensional. For composite systems, we define the tensor product of  $\mathcal{H}_A$  and  $\mathcal{H}_B$  as  $\mathcal{H}_{AB} := \mathcal{H}_A \otimes \mathcal{H}_B$ . We denote  $\mathcal{P}(\mathcal{H})$  as the set of positive semi-definite operators on  $\mathcal{H}$  and  $\mathcal{S}(\mathcal{H})$  as the set of normalized states on  $\mathcal{H}$ , i.e.,  $\mathcal{S}(\mathcal{H}) = \{\rho \in \mathcal{P}(\mathcal{H}) : \text{tr}(\rho) = 1\}$ . Furthermore, for a composite state  $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ , the reduced states of system A and system B are given by  $\rho_A = \text{tr}_B(\rho_{AB})$  and  $\rho_B = \text{tr}_A(\rho_{AB})$ , respectively. A positive operator valued measure (POVM) is denoted by  $\mathbb{M} := \{M_x\}_x$  where  $\sum_x M_x = \mathbb{1}$ . For any POVM, we may view it as a projective measurement by introducing an ancillary system, thus for any POVM with binary outcomes, we may write it as an observable  $O = \sum_{x \in \{0,1\}} (-1)^x M_x$ , such that  $\sum_{x \in \{0,1\}} M_x = \mathbb{1}$ . We also use  $\bar{x} := (x_1, x_2, \dots, x_n)$  to represent the concatenations of elements and  $[n]$  to denote  $\{1, 2, \dots, n\}$ .

### 2. Assumptions required for the security proof

Prior to stating the security proof, it is instructive to elucidate the assumptions which are necessary for the security proof. In particular, we define the minimal amount of knowledge and resources that are required for the security proof. The assumptions are detailed in the following:

- A1 Trusted local sources of randomness.** Alice (also Bob) has access to a trusted source that produces a random and secure bit value upon each use. Furthermore, we assume the source is unlimited, that is, Alice can use it as much as she wants, however the protocol only requires an amount of randomness linear in the number of quantum states generated.



**A2 An authenticated but otherwise insecure classical channel.** Generally, this assumption is satisfied if Alice and Bob share an initial short secret key [37, 38]. Note that the security analysis of such authentication schemes were recently extended to the universally composable framework by Ref [39], which allows one to compose the error of the authentication scheme with the errors of the protocol, giving an overall error on the security.

**A3 No information is allowed to leave the laboratories unless the protocol prescribes it.** This assumption is paramount to any cryptographic protocol, and roughly speaking, it states that information generated by the legitimate users is appropriately controlled. More concretely, we assume the followings

- (a) *Communication lines.*— The only two communication lines leaving the laboratory are the classical and the quantum channel. Furthermore, the classical channel is controlled, i.e., only the information required by the protocol is sent.
- (b) *Communication between devices.*— There should be no unauthorized communication between any devices in the laboratory, in particular from the measurement devices to the source device.

**A4 Trusted classical operations.** Classical operations like authentication, error correction, error verification, privacy amplification, etc must be trusted, i.e., we know that the operations have ideal functionality and are independent of the adversary.

**A5 \*Measurement and source devices have no internal memories.** This implies each use of the device is independent of the previous uses. For example, for  $N$  uses of a source device and a measurement that produces a bit string  $\bar{x} := (x_1, x_2, \dots, x_n)$ , we have

$$\rho^N = \bigotimes_{i=1}^N \rho^i, \quad M_{\bar{x}} = \bigotimes_i M_{x_i}^i$$

where  $M_{\bar{x}}$  is the POVM element corresponding to the outcome  $\bar{x}$ .

**A6 \*The marginal states of Alice and Bob are independent of whether Charlie outputs a pass or fail.** Mathematically, assumption A6 corresponds to the following: let  $\rho_{AC}$  be the bipartite state generated by Alice and let  $\rho_A = \text{Tr}_C(\rho_{AC})$  be the marginal state of Alice, then assumption A6 is satisfied by the identity

$$\rho_{A|\text{pass}} = \rho_{A|\text{fail}}$$

where  $\rho_{A|\text{pass}}$  and  $\rho_{A|\text{fail}}$  are the marginal states of Alice conditioned on Charlie outputting pass and fail, respectively.

In general, assumptions A1, A2, A3 and A4 are necessary for all cryptography protocols, be it quantum or classical. However, to prove the security of the protocol, we require additional assumptions A5 and A6 (denoted by the superscript  $\star$  in the list). The former is required to simplify the problem at hand, that is, by restricting our analysis to devices that have no internal memories, the statistical analysis required for the security proof can be restricted to independent systems. The latter is required to prevent the adversary from establishing correlations with Alice's measurement and source devices. In particular, it is to prevent the adversary from post-selecting measurement outcomes (or the raw key<sup>5</sup>) that are favorable to her. For example, Charlie (in the worst case, an accomplice of the adversary) can choose to output pass if only insecure states are produced by the source of Alice (also Bob), otherwise he outputs fail. This predicament appears to be unavoidable when one considers a general source device—where the physics is unknown—for quantum cryptography, i.e., the systems which leave the laboratory might carry valuable information for the adversary, e.g., such devices can be motivated to leak out crucial past information [40]. Although we do not consider completely general devices (or devices with adversarial motives), some form of guarantee on the observed statistics is still required. That is to say, we either compute the amount of information the adversary has on the raw key in the worst case scenario (see section 7) or adopt an ad-hoc assumption<sup>6</sup> like assumption A6 which prevents Charlie from establishing correlations with Alice's and Bob's measurement and source devices. For simplicity reasons, we adopt assumption A6.

<sup>5</sup> The raw key is defined as the measurement outcomes used to obtain the secret key.

<sup>6</sup> In fact, introducing ad-hoc assumptions is not new, e.g., Ref [41] made the assumption that the dimension of the source device is fixed

### 3. Technical lemmas

**Lemma 1** (Jordan's lemma [25, 42, 43]). *Let  $O$  and  $O'$  be observables with eigenvalues  $\pm 1$  on Hilbert space  $\mathcal{H}$ . Then there exists a partition of the Hilbert space,  $\mathcal{H} = \bigoplus_i \mathcal{H}_i$ , such that*

$$O = \bigoplus_i O_i \quad \text{and} \quad O' = \bigoplus_i O'_i$$

where  $\mathcal{H}_i$  satisfies  $\dim(\mathcal{H}_i) \leq 2$  for all  $i$ .

**Lemma 2** (Chernoff-Hoeffding [44]). *Let  $X := \frac{1}{n} \sum_i X_i$  be the average of  $n$  independent random variables  $X_1, X_2, \dots, X_n$  with values in  $[0, 1]$ , and let  $\mu := \mathbb{E}[X] = \frac{1}{n} \sum_i \mathbb{E}[X_i]$  denote the expected value of  $X$ . Then, for any  $\delta > 0$ ,*

$$\Pr[X - \mu \geq \delta] \leq \exp(-2\delta^2 n).$$

**Lemma 3** (Serfling [45]). *Let  $\{x_1, \dots, x_n\}$  be a list of (not necessarily distinct) values in  $[a, b]$  with average  $\mu := \frac{1}{n} \sum_i x_i$ . Let the random variables  $X_1, X_2, \dots, X_t$  be obtained by sampling  $t$  random entries from this list without replacement. Then, for any  $\delta > 0$ , the random variable  $X := \frac{1}{t} \sum_i X_i$  satisfies*

$$\Pr[X - \mu \geq \delta] \leq \exp\left(\frac{-2\delta^2 tn}{(n-t+1)(b-a)}\right).$$

**Lemma 4** (Generalized UCR for commuting measurements [13]). *Let  $\varepsilon > 0, \bar{\varepsilon} \geq 0$  and  $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ . Moreover let  $\mathbf{M} = \{M_x\}$ ,  $\mathbf{N} = \{N_z\}$  be POVMs on  $\mathcal{H}_A$ , and  $\mathbf{K} = \{P_k\}$  a projective measurement on  $\mathcal{H}_A$  that commutes with both  $\mathbf{M}$  and  $\mathbf{N}$ . Then the post-measurement states*

$$\rho_{XB} = \sum_x |x\rangle\langle x| \otimes \text{tr}_{AC}(\sqrt{M_x} \rho_{ABC} \sqrt{M_x}), \quad \rho_{ZC} = \sum_z |z\rangle\langle z| \otimes \text{tr}_{AB}(\sqrt{N_z} \rho_{ABC} \sqrt{N_z}),$$

satisfy

$$H_{\min}^{2\varepsilon+\bar{\varepsilon}}(X|B)_\rho + H_{\max}^\varepsilon(Z|C)_\rho \geq \log \frac{1}{c^*(\rho_A, \mathbf{M}, \mathbf{N})} - \log \frac{2}{\bar{\varepsilon}^2}, \quad (\text{S1})$$

where the effective overlap is defined as

$$c^*(\rho_A, \mathbf{M}, \mathbf{N}) := \min_{\mathbf{K}} \left\{ \sum_k \text{tr}(P_k \rho) \max_x \|P_k \sum_z N_z M_x N_z\|_\infty \right\} \quad (\text{S2})$$

Note that (S1) is a statement about the entropies of the post-measurement states  $\rho_{XB}$  and  $\rho_{ZC}$ , thus it also holds for any measurements that lead to the same post-measurement states. Accordingly, one may also consider the projective purifications  $\mathbf{M}'$  and  $\mathbf{N}'$  of  $\mathbf{M}$  and  $\mathbf{N}$ , applied to  $\rho_A \otimes |\phi\rangle\langle\phi|$ , where  $|\phi\rangle$  is a pure state of an ancilla system. Since both measurement setups  $\{\rho, \mathbf{M}, \mathbf{N}\}$  and  $\{\rho_A \otimes |\phi\rangle\langle\phi|, \mathbf{M}', \mathbf{N}'\}$  give the same post-measurement states, the R.H.S of (S1) holds for both  $c^*(\rho_A, \mathbf{M}, \mathbf{N})$  and  $c^*(\rho_A \otimes |\phi\rangle\langle\phi|, \mathbf{M}', \mathbf{N}')$ . We can thus restrict our considerations to projective measurements.

In the protocol considered, Alice performs independent binary measurements —  $\mathbf{M}_i = \{M_x^i\}_{x \in \{0,1\}}$  and  $\mathbf{N}_i = \{N_z^i\}_{z \in \{0,1\}}$  — on each subsystem  $i$ . We can reduce (S2) to operations on each subsystem, if we choose  $\mathbf{K} = \{P_{\bar{k}}\}$  to also be in product form, i.e.,  $P_{\bar{k}} = \bigotimes_i P_{k_i}^i$ , where  $\bar{k}$  is a string of (not necessarily binary) letters  $k_i \in \mathcal{K}$ . Then plugging this,  $M_{\bar{x}} = \bigotimes_i M_{x_i}^i$  and  $N_{\bar{z}} = \bigotimes_i N_{z_i}^i$  in the norm from (S2), we get

$$\|P_{\bar{k}} \sum_{\bar{z}} N_{\bar{z}} M_{\bar{x}} N_{\bar{z}}\|_\infty = \left\| \sum_{z_1, z_2, \dots} \bigotimes_i P_{k_i}^i N_{z_i}^i M_{x_i}^i N_{z_i}^i \right\|_\infty = \prod_i \|P_{k_i}^i \sum_{z_i} N_{z_i}^i M_{x_i}^i N_{z_i}^i\|_\infty.$$

Putting this in (S2) with  $\rho = \bigotimes_i \rho^i$ ,  $p_k^i := \text{tr}(P_k^i \rho^i)$ , and dropping the subscript  $i$  when possible, we obtain,

$$\begin{aligned} c^*(\rho_A, \mathbf{M}, \mathbf{N}) &\leq \sum_{k_1, k_2, \dots} \prod_i p_{k_i}^i \max_x \|P_{k_i}^i \sum_z N_z^i M_x^i N_z^i\|_\infty \\ &= \prod_i \sum_k p_k^i \max_x \|P_k^i \sum_z N_z^i M_x^i N_z^i\|_\infty =: \prod_i c^{*,i}. \end{aligned} \quad (\text{S3})$$

In the following we will refer to

$$c_k^i := \max_x \|P_k^i \sum_z N_z^i M_x^i N_z^i\|_\infty \quad (\text{S4})$$

as the overlap of the measurements  $\{M_x^i\}_x$  and  $\{N_z^i\}_z$ .

#### 4. An upper bound on the effective overlap with the CHSH value

In this subsection, we first introduce the notion of CHSH operator [46] and then prove the relation between the CHSH test and the effective overlap (S4).

In the CHSH test, two space-like separated systems share a bipartite state  $\rho$  and each system has two measurements. More specifically, system A has POVMs  $\{M_0^0, M_1^0\}$  and  $\{M_0^1, M_1^1\}$  and system T has POVMs  $\{T_0^0, T_1^0\}$  and  $\{T_0^1, T_1^1\}$ . Since for any POVM there is a (unitary and) projective measurement on a larger Hilbert space that has the same statistics, we can restrict our considerations to projective measurements. Then, we may write the POVMs as observables with  $\pm 1$  outcomes, i.e., at the site of the first system, the two observables are  $O_A^0 := \sum_{s=0}^1 (-1)^s M_s^0$  and  $O_A^1 := \sum_{s=0}^1 (-1)^s M_s^1$ . Furthermore, the measurements are chosen uniformly at random. As such, the CHSH value is given by  $S(\rho, \beta) := \text{Tr}(\rho\beta)$  where the CHSH operator is defined as

$$\beta(O_A^0, O_A^1, O_T^0, O_T^1) := \sum_{u,v} (-1)^{u \wedge v} O_A^u \otimes O_T^v \quad (\text{S5})$$

where  $u, v$  and  $s, t$  are the inputs and outputs, respectively. The maximization of  $S(\rho, \beta)$  over the set of density operators for a fixed  $\beta$  is defined by  $S_{\max}(\beta)$ . Moreover, the CHSH operator can be decomposed into a direct sum of two-qubits subspaces via Lemma 1. Mathematically, we may write  $O_A^0 = \sum_k P_k O_A^0 P_k$  and  $O_A^1 = \sum_k P_k O_A^1 P_k$  where  $\{P_k\}_k$  is a set of projectors such that  $\dim(P_k) = 2 \forall k$ . Note that in Lemma 1, one may select a partition of the Hilbert space such that each block partition has dimension two. This allows one to decompose the general CHSH operator into direct sums of qubits CHSH operators. Likewise, for the measurements of Bob,  $O_B^0 = \sum_r Q_r O_T^0 Q_r$  and  $O_B^1 = \sum_r Q_r O_T^1 Q_r$ . For all  $k$ ,  $P_k O_A^0 P_k$  and  $P_k O_A^1 P_k$  can be written in terms of Pauli operators,

$$P_k O_A^0 P_k = \vec{m}_k \cdot \Gamma_k \quad \text{and} \quad P_k O_A^1 P_k = \vec{n}_k \cdot \Gamma_k, \quad (\text{S6})$$

where  $\vec{m}_k$  and  $\vec{n}_k$  are unit vectors in  $\mathbb{R}_k^3$  and  $\Gamma_k$  is the Pauli vector. Combining (S5) and (S6) yields

$$\beta = \bigoplus_{k,r} \beta_{k,r} \quad \text{where} \quad \beta_{k,r} \in \mathbb{C}_k^2 \otimes \mathbb{C}_r^2 \quad (\text{S7})$$

and it can be verified that

$$S(\rho, \beta) = \sum_{k,r} \lambda_{k,r} S_{k,r} \quad (\text{S8})$$

where

$$\lambda_{k,r} := \text{Tr}(P_k \otimes Q_r \rho) \quad (\text{S9})$$

$$S_{k,r} := \text{Tr}(\rho_{k,r} \beta_{k,r}) \quad (\text{S10})$$

Whenever the context is clear, we write  $S = S(\rho, \beta)$  and  $S_{\max} = S_{\max}(\beta)$ .

In the following analysis, we consider only one subsystem, the superscript  $i$  is omitted, i.e., we use  $c^* = \sum_k p_k c_k$  instead.

**Lemma 5.** *Let  $\{O_A^x\}_{x \in \{0,1\}}$  and  $\{O_T^y\}_{y \in \{0,1\}}$  be observables with eigenvalues  $\pm 1$  on  $\mathcal{H}_A$  and  $\mathcal{H}_T$  respectively and let  $\beta = \sum_{x,y} (-1)^{x \wedge y} O_A^x \otimes O_T^y$  be the CHSH operator. Then for any  $\rho \in \mathcal{S}(\mathcal{H}_{AT})$ , the effective overlap  $c^*$  is related to the CHSH value  $S = \text{Tr}(\rho\beta)$  by*

$$c^* \leq \frac{1}{2} + \frac{S}{8} \sqrt{8 - S^2} \quad (\text{S11})$$



*Proof.* Using (S6), let the relative angle between  $\vec{m}_k$  and  $\vec{n}_k$  be  $\theta_k \in [0, \pi/2]$  for all  $k$ , i.e.,  $\vec{m}_k \cdot \vec{n}_k = \cos(\theta_k)$ . Furthermore, we can express  $\vec{m}_k \cdot \Gamma_k$  and  $\vec{n}_k \cdot \Gamma_k$  in terms of rank-1 projectors. Formally, we have  $\vec{m}_k \cdot \Gamma_k = |\vec{m}_k\rangle\langle\vec{m}_k| - |-\vec{m}_k\rangle\langle-\vec{m}_k|$  and similarly for  $\vec{n}_k \cdot \Gamma_k$ . Plugging these into (S4),

$$c_k = \max_{i,j \in \{0,1\}} |((-1)^i \vec{m}_k | (-1)^j \vec{n}_k)|^2 = \frac{1 + \cos \theta_k}{2} \quad (\text{S12})$$

Next, we want to relate  $c_k$  to the CHSH value. Using the result of Seevinck and Uffink [31], for all  $r$ , (S10) satisfies

$$S_{k,r} \leq 2\sqrt{1 + \sin(\theta_k) \sin(\theta_r)} \quad (\text{S13})$$

where  $\sin(\theta_k)$  and  $\sin(\theta_r)$  quantify the commutativity of Alice's  $k$ th and system T's  $r$ th measurements, respectively. From (S12) and (S13) we obtain for all  $r$ ,

$$c_k \leq \frac{1}{2} + \frac{S_{k,r}}{8} \sqrt{8 - S_{k,r}^2},$$

where we use the fact that the right hand side is a monotonic decreasing function. Finally, we get

$$c^* = \sum_k p_k c_k = \sum_{k,r} \lambda_{k,r} c_k \leq \frac{1}{2} + \frac{S}{8} \sqrt{8 - S^2},$$

and the inequality is given by the Jensen's inequality and (S8). □

## 5. Statistics

We recall in the protocol definition, after the sifting step, Alice and Bob have sets  $\mathcal{X}, \mathcal{Z}, \mathcal{J}, \mathcal{J}'$  which correspond to key generation, channel error estimation, Alice's CHSH test and Bob's CHSH test. We need to estimate the average overlap of set  $\mathcal{X}$  given the observed CHSH value evaluated on sets  $\mathcal{J}$  and  $\mathcal{J}'$ . To do that, we need the following two statistical statements: the first statement (Lemma 6) gives a bound on the probability that the observed CHSH value is larger than the expected CHSH value and the second statement (Lemma 7) gives a bound on the probability that the average of the values  $c^{*,i}$  for  $i \in \mathcal{X}$  (used to generate the key) is larger than the average of the values  $c^{*,i}$  for  $i \in \mathcal{J}$  (used for the CHSH test).

**Lemma 6.** *Let  $S_{\mathcal{J}}$  be the average CHSH value on  $j$  independent systems, and  $S_{\text{test}}$  the observed CHSH on these systems. Then*

$$\Pr \left[ S_{\text{test}} - S_{\mathcal{J}} \geq \sqrt{\frac{32}{j} \ln \frac{1}{\varepsilon}} \right] \leq \varepsilon.$$

*Proof.* We define the random variable

$$Y_i := \begin{cases} 1 & \text{if } s_i \oplus t_i = u_i \wedge u_i, \\ 0 & \text{otherwise,} \end{cases}$$

where  $u_i, v_i, s_i, t_i$  are the inputs and outputs, respectively of the measurements on system  $i$ , and  $Y_{\mathcal{J}} := \frac{1}{j} \sum_{i \in \mathcal{J}} Y_i$ . It is easy to see that  $S_i = 8 \mathbb{E}[Y_i] - 4$ ,  $S_{\mathcal{J}} = 8 \mathbb{E}[Y_{\mathcal{J}}] - 4$  and  $S_{\text{test}} = Y_{\mathcal{J}}$ . The proof is then immediate from Lemma 2. □

In the main text, we made the assumption A6 that the marginal states of Alice and Bob are independent of Charlie's measurement outcome. This implies that Charlie's measurement does not distinguish between setups with small and large  $c^{*,i}$ . As a consequence, the average  $c^*$  value on  $\mathcal{X}$  can be estimated with the average  $c^*$  value on  $\mathcal{J}$ .

**Lemma 7.** *Let  $c_{\mathcal{X}}^* := \frac{1}{m_x} \sum_{i \in \mathcal{X}} c^{*,i}$  be the average  $c^*$  value on the set  $\mathcal{X}$  used to generate the secret key, and  $c_{\mathcal{J}}^* := \frac{1}{j} \sum_{i \in \mathcal{J}} c^{*,i}$  be the average  $c^*$  value on the set  $\mathcal{J}$  selected for the CHSH test. Then*

$$\Pr \left[ c_{\mathcal{X}}^* - c_{\mathcal{J}}^* \geq \sqrt{\frac{(m_x + j)(j + 1)}{2m_x j^2} \ln \frac{1}{\varepsilon}} \right] \leq \varepsilon.$$

*Proof.* Let  $\mu := \frac{1}{m_x+j} \sum_{i \in \mathcal{X} \cup \mathcal{J}} c^{*,i}$ . Since the adversary cannot distinguish between systems with small and large  $c^{*,i}$ ,  $\mathcal{J}$  can be seen as a random subset of  $\mathcal{X} \cup \mathcal{J}$ . Hence we can apply Lemma 3, from which we get

$$\Pr[c_{\mathcal{X}}^* - \mu \geq \delta] \leq \exp \frac{-2\delta^2 m_x(m_x + j)}{j + 1}.$$

Using  $\mu = \frac{m_x}{m_x+j} c_{\mathcal{X}}^* + \frac{j}{m_x+j} c_{\mathcal{J}}^*$  we finish the proof.  $\square$

## 6. Secrecy of the protocol

With the relevant results in hand, we are ready to prove our main result which follows roughly the same line of argument as Ref [10]. The main differences are the use of a more general smooth entropic uncertainty relation (Lemma 4) to bound the error on the secrecy, and of the CHSH test to bound the effective overlap of the measurement operators and states used by the uncertainty relation (Lemma 5). Since the players can only sample the CHSH violation, we use Lemma 7 to bound the distance between this estimate and the expected effective overlap of the key set. The correctness of the protocol are evaluated in exactly the same way as in Ref [10], so we refer to that work for the corresponding bounds and theorems. We only prove the secrecy of the protocol here.

Contrary to most QKD protocols, the protocol adopts a tripartite model where Charlie is supposed to establish entanglement between Alice and Bob. Thus in our picture, we can view Charlie as an accomplice of the adversary and evaluate the secrecy on the overall state conditioned on the events whereby Charlie outputs a pass.

We briefly recall the main parameters of the protocol, which are detailed in the protocol definition given in the main text. Conditioned on the successful operation of Charlie (the events whereby Charlie outputs a pass), Alice and Bob generate systems until at least  $m_x$  of them have been measured by both of them in the basis X,  $m_z$  have been measured in the basis Z, and  $j$  have been chosen for both CHSH tests. The tolerated error rate and the CHSH value are  $Q_{\text{tol}}$  and  $S_{\text{tol}}$ , respectively. Furthermore, we assume that our information reconciliation scheme leaks at most  $\text{leak}_{\text{EC}} + \lceil \log(1/\varepsilon_{\text{cor}}) \rceil$ -bits of information, where an error correction scheme which leaks at most  $\text{leak}_{\text{EC}}$ -bits of information is applied, then an error verification scheme which leaks  $\lceil \log(1/\varepsilon_{\text{cor}}) \rceil$ -bits of information is applied. If the error verification fails, they abort the protocol. Note that Alice and Bob should check who has the higher CHSH value, then the information reconciliation scheme is implemented from that party's point of view. In the following, we assume that Alice always has the higher CHSH value, i.e., Bob is supposed to reconstruct the key of Alice.

**Theorem 1.** *The protocol is  $\varepsilon_{\text{sec}}$ -secret if for some  $\varepsilon_Q, \varepsilon_{\text{UCR}}, \varepsilon_{\text{PA}}, \varepsilon_{c^*}, \varepsilon_{\text{CHSH}} > 0$  such that  $2\varepsilon_Q + \varepsilon_{\text{UCR}} + \varepsilon_{\text{PA}} + \varepsilon_{c^*} + \varepsilon_{\text{CHSH}} \leq \varepsilon_{\text{sec}}$ , the final secret key length  $\ell$  satisfies*

$$\ell \leq m_x \left( 1 - \log \left( 1 + \frac{\hat{S}_{\text{tol}}}{4} \sqrt{8 - \hat{S}_{\text{tol}}^2} + \zeta(\varepsilon_{c^*}) \right) - h(\hat{Q}_{\text{tol}}) \right) - \text{leak}_{\text{EC}} - \log \frac{2}{\varepsilon_{\text{UCR}}^2} - \log \frac{2}{\varepsilon_{\text{PA}}^2 \varepsilon_{\text{cor}}}, \quad (\text{S14})$$

where  $\hat{S}_{\text{tol}} := S_{\text{tol}} - \xi(\varepsilon_{\text{CHSH}})$  and  $\hat{Q}_{\text{tol}} := Q_{\text{tol}} + \mu(\varepsilon_Q)$  with the statistical fluctuations given by

$$\xi(\varepsilon_{\text{CHSH}}) := \sqrt{\frac{32}{j} \ln \frac{1}{\varepsilon_{\text{CHSH}}}}, \quad \zeta(\varepsilon_{c^*}) := \sqrt{\frac{2(m_x + j)(j + 1)}{m_x j^2} \ln \frac{1}{\varepsilon_{c^*}}}, \quad \text{and} \quad \mu(\varepsilon_Q) := \sqrt{\frac{(m_x + m_z)(m_z + 1)}{m_x m_z^2} \ln \frac{1}{\varepsilon_Q}}.$$

*Proof.* If one of the tests  $Q_{\text{test}} \leq Q_{\text{tol}}$  and  $S_{\text{test}} \geq S_{\text{tol}}$  fails, then the protocol aborts, and the secrecy error is trivially zero. Conditioned on passing these tests, let  $X$  be the raw key of length  $m_x$  that Alice gets from the set  $\mathcal{X}$ , and let  $E$  denote the adversary's information obtained by eavesdropping on the quantum channel. After listening to the error correction and hash value, Eve has a new system  $E'$ . Using  $\text{leak}_{\text{EC}} + \lceil \log(1/\varepsilon_{\text{cor}}) \rceil \leq \text{leak}_{\text{EC}} + \log(2/\varepsilon_{\text{cor}})$  (the number bits used for error correction and error verification) and using chain rules for smooth entropies [13] we can bound the min-entropy of the raw key  $X$  given  $E'$

$$H_{\min}^{2\varepsilon + \varepsilon_{\text{UCR}}}(X|E') \geq H_{\min}^{2\varepsilon + \varepsilon_{\text{UCR}}}(X|E) - \text{leak}_{\text{EC}} - \log \frac{2}{\varepsilon_{\text{cor}}}.$$

From the entropic uncertainty relation (Lemma 4), we further get

$$H_{\min}^{2\varepsilon + \varepsilon_{\text{UCR}}}(X|E) \geq \log \frac{1}{c^*} - H_{\max}^{\varepsilon}(Z|B) - \log \frac{2}{\varepsilon_{\text{UCR}}^2},$$

where  $Z$  can be seen as the outcome Alice would have gotten if she had measured the same systems in the corresponding basis  $Z$ , and  $B$  is Bob's system in this case (before measurement).

The max-entropy of the alternative measurement is then bounded by the error rate sampled on the  $m_z$  systems  $\mathcal{Z}$  [10]:

$$H_{\max}^{\varepsilon}(Z|B) \leq m_x h(Q_{\text{tol}} + \mu(\varepsilon_Q)),$$

where  $h(x) := -x \log_2 x - (1-x) \log_2 (1-x)$  is the binary entropy function of  $x$ ,  $\varepsilon = \varepsilon_Q / \sqrt{p_{\text{pass}}}$  and  $p_{\text{pass}} := \Pr[Q_{\text{test}} \leq Q_{\text{tol}} \wedge S_{\text{test}} \geq S_{\text{tol}}]$  is the probability of the event “pass”, namely that this round of QKD passes the statistical tests.

Next, we bound  $c^*$  in terms of the measured CHSH value  $S_{\text{test}}$ . We first use the arithmetic-geometric means inequality, from which we get

$$c^* \leq \prod_{i \in \mathcal{X}} c^{*,i} \leq \left( \sum_{i \in \mathcal{X}} \frac{c^{*,i}}{m_x} \right)^{m_x} = (c_{\mathcal{X}})^{m_x}.$$

Since the adversary cannot distinguish between systems with small and large  $c^{*,i}$ ,  $c_{\mathcal{X}}^*$  can be seen as a random sample from  $c_{\mathcal{J} \cup \mathcal{X}}^*$ , where  $\mathcal{J}$  is the set of systems used for the CHSH test. From Lemma 7 we get  $\Pr[c_{\mathcal{X}} - c_{\mathcal{J}} \geq \zeta(\varepsilon_{c^*})/2] \leq \varepsilon_{c^*}$ , hence

$$\varepsilon' := \Pr \left[ c_{\mathcal{X}} - c_{\mathcal{J}} \geq \frac{\zeta(\varepsilon_{c^*})}{2} \middle| \text{“pass”} \right] \leq \frac{\varepsilon_{c^*}}{p_{\text{pass}}}.$$

Lemma 5 can now be used together with Jensen's inequality, so with probability at least  $1 - \varepsilon'$ ,

$$c_{\mathcal{X}} \leq \frac{1}{2} \left( 1 + \frac{S_{\mathcal{J}}}{4} \sqrt{8 - S_{\mathcal{J}}^2} + \zeta(\varepsilon_{c^*}) \right).$$

We still need to take into account that we only have an approximation for the CHSH value of the systems in  $\mathcal{J}$ . From Lemma 6 we get that

$$\varepsilon'' := \Pr \left[ S_{\mathcal{J}} \leq \hat{S}_{\text{test}} \middle| \text{“pass”} \right] \leq \frac{\varepsilon_{\text{CHSH}}}{p_{\text{pass}}}.$$

Finally, the bound on the error of privacy amplification by universal hashing [32, 47, 48] says that the error is less than  $2\varepsilon + \varepsilon_{\text{UCR}} + \varepsilon_{\text{PA}}$  as long as

$$\ell \leq H_{\min}^{2\varepsilon + \varepsilon_{\text{UCR}}}(X|E') - 2 \log \frac{1}{\varepsilon_{\text{PA}}}.$$

Putting all the above equations together we get (S14), with a total error conditioned on passing the tests  $Q_{\text{test}} \leq Q_{\text{tol}}$  and  $S_{\text{test}} \geq S_{\text{tol}}$  of at most  $2\varepsilon + \varepsilon_{\text{UCR}} + \varepsilon_{\text{PA}} + \varepsilon' + \varepsilon''$ . If we remove this conditioning, the error is then

$$p_{\text{pass}}(2\varepsilon + \varepsilon_{\text{UCR}} + \varepsilon_{\text{PA}} + \varepsilon' + \varepsilon'') \leq 2\varepsilon_Q + \varepsilon_{\text{UCR}} + \varepsilon_{\text{PA}} + \varepsilon_{c^*} + \varepsilon_{\text{CHSH}}. \quad \square$$

## 7. A method to remove assumption A6

In our security proof, we require that the adversary cannot distinguish between rounds with small and large effective overlap  $c^{*,i}$ . Here, we sketch an approach to remove assumption A6 and show that the secret key rate of the protocol in the most perilous scenario is generally dependent on the channel loss; only in the limiting scenario—where the observed CHSH value is maximal—it is independent of the channel loss.

To show the above, we only need to modify Lemma 7. First, let  $c_{\mathcal{X}'}^* := \sum_{i \in \mathcal{X}'} c^{*,i}$  where the set  $\mathcal{X}'$  denotes the systems whereby Alice and Bob measure in the  $X$  basis. Clearly, we have  $\mathcal{X} \subseteq \mathcal{X}'$  with equality only if Charlie's operation has unit efficiency, i.e.,  $\eta_{\text{eff}} = 1$ . Next, we consider  $\{c^{*,i}\}_{i \in \mathcal{X}'}$  in decreasing order, that is,  $c^{*,1} \geq c^{*,2} \geq \dots \geq c^{*,|\mathcal{X}'|}$ . Then the average overlap of  $\mathcal{X}'$  is decomposed as

$$c_{\mathcal{X}'}^* = \frac{m_x}{|\mathcal{X}'|} \sum_{i=1}^{m_x} \frac{c^{*,i}}{m_x} + \sum_{j=m_x+1}^{|\mathcal{X}'|} \frac{c^{*,i}}{|\mathcal{X}'|} \geq \frac{m_x}{|\mathcal{X}'|} \left( c_{\mathcal{X}}^* - \frac{1}{2} \right) + \frac{1}{2}$$

where we assume that the adversary selects the systems for  $\mathcal{X}$  and the inequality is given by using  $c^{*,i} \geq 1/2$ . As such, the average overlap of  $\mathcal{X}$  satisfies

$$c_{\mathcal{X}}^* \leq \frac{1}{2} + \frac{1}{\eta_{\text{eff}}} \left( c_{\mathcal{X}'}^* - \frac{1}{2} \right) \quad \text{where} \quad \eta_{\text{eff}} := \frac{m_x}{|\mathcal{X}'|} \quad (\text{S15})$$

Note that in practical scenarios,  $\eta_{\text{eff}}$  can be identified with the efficiency of the Bell state measurement and the channel loss. Finally, since  $c_{\mathcal{X}'}^*$  can be estimated from  $c_{\mathcal{J}}^*$  via Lemma 7, the security proof follows from the previous sub-section. In the asymptotic limit, using (S15) and Lemma 5, it is easy to verify that our protocol is secure as long as the condition  $4\eta_{\text{eff}} > S_{\text{test}} \sqrt{8 - S_{\text{test}}^2}$  holds. Accordingly, the distance between Alice and Bob is limited by the strength of the CHSH violation.

- 
- [1] Gisin *et al.*, *Quantum cryptography.*, Rev. Mod. Phys., **74**, 145–195, (2002); Scarani *et al.*, *The security of practical quantum key distribution*, Rev. Mod. Phys., **81**, 1301–1350, (2009).
  - [2] C.-H. F. Fung *et al.*, *Phase-remapping attack in practical quantum-key-distribution systems*, Phys. Rev. A. **75**, 032314 (2007); A. Lamas-Linares and C. Kurtsiefer, *Breaking a quantum key distribution system through a timing side channel*, Opt. Express **15**, 9388 (2007); S. Nauerth *et al.*, *Information leakage via side channels in freespace BB84 quantum cryptography*, New J. Phys. **11**, 065001 (2009); L. Lydersen *et al.*, *Hacking commercial quantum cryptography systems by tailored bright illumination*, Nature. Photon. **4**, 686, (2010); H. Weier *et al.*, *Quantum Eavesdropping without Interception: An Attack Exploiting the Dead Time of Single Photon Detectors*, New J. Phys. **13**, 073024 (2011); I. Gerhardt *et al.*, *Full-field implementation of a perfect eavesdropper on a quantum cryptography system.*, Nature Commun. **2**, 349 (2011).
  - [3] D. Gottesman *et al.*, *Security of quantum key distribution with imperfect devices*, Quantum Inf. Comput. **5**, 325 (2004).
  - [4] M. Hayashi, *Practical evaluation of security for quantum key distribution*, Phys. Rev. A. **74**, 22307 (2006).
  - [5] H. Inamori, N. Lütkenhaus, and D. Mayers, *Unconditional security of practical quantum key distribution*, Eur. Phys. J. D **41**, 599 (2007).
  - [6] V. Scarani and R. Renner, *Quantum Cryptography with Finite Resources: Unconditional Security Bound for Discrete-Variable Protocols with One-Way Postprocessing*, Phys. Rev. Lett. **100**, 200501 (2008).
  - [7] L. Sheridan, T. P. Le, and V. Scarani, *Finite-key security against coherent attacks in quantum key distribution*, New J. Phys. **12**, 123019 (2010).
  - [8] S. Bratzik *et al.*, *Min-entropy and quantum key distribution: non-zero key rates for "small" numbers of signals*, Phys. Rev. A **83**, 022330 (2011).
  - [9] M. Hayashi and T. Tsurumaru, *Concise and Tight Security Analysis of the Bennett-Brassard 1984 Protocol with Finite Key Lengths*, arXiv:1107.0589, (2011).
  - [10] M. Tomamichel *et al.*, *Tight finite-key analysis for quantum cryptography*, Nature Commun. **3**, 634 (2012).
  - [11] M. Berta *et al.*, *The uncertainty principle in the presence of quantum memory*, Nature Phys. **6** 659, (2010).
  - [12] M. Tomamichel and R. Renner, *Uncertainty relation for smooth entropies*, Phys. Rev. Lett. **106**, 110506 (2010).
  - [13] M. Tomamichel, *A Framework for Non-Asymptotic Quantum Information Theory*, Diss. ETH. No. 20213, arXiv:1203.2142, (2012).
  - [14] E. Hänggi and M. Tomamichel, *The Link between Uncertainty Relations and Non-Locality*, arXiv:1108.5349, (2011).
  - [15] J F Clauser *et al.*, *Proposed Experiment to Test Local Hidden-Variable Theories*, Phys. Rev. Lett. **23**, 880 (1969).
  - [16] J. S. Bell, *On the Einstein Podolsky Rosen paradox*, Physics. **1**, 195 (1964).
  - [17] P. Pearle, *Hidden-Variable Example Based upon Data Rejection*, Phys. Rev. D. **2**, 1418 (1970).
  - [18] E. Biham, B. Huttner and T. Mor, *Quantum cryptographic network based on quantum memories*, Phys. Rev. A. **54**, 2651 (1996).
  - [19] H. Inamori, *Security of Practical Time-Reversed EPR Quantum Key Distribution*, Algorithmica **34**, 340 (2002).
  - [20] S. L. Braunstein and S. Pirandola *Side-Channel-Free Quantum Key Distribution*, Phys. Rev. Lett. **108**, 130503 (2012).
  - [21] H.-K. Lo, M. Curty and B. Qi, *Measurement-Device-Independent Quantum Key Distribution*, Phys. Rev. Lett. **108**, 130503 (2012).
  - [22] C. H. Bennett *et al.*, *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*, Phys. Rev. Lett. **70**, 1895 (1993).
  - [23] C. H. Bennett and G. Brassard, *Quantum cryptography: Public key distribution and coin tossing*, in Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing (IEEE Press, Bangalore, India New York, 1984), p. 175.
  - [24] A. K. Ekert, *Quantum cryptography based on Bell's theorem*, Phys. Rev. Lett. **67**, 661 (1991).
  - [25] S. Pironio *et al.*, *Device-Independent Quantum Key Distribution Secure Against Collective Attacks*, New J. Phys. **11**, 045021 (2009).
  - [26] M. McKague, *Device independent quantum key distribution secure against coherent attacks with memoryless measurement devices*, New J. Phys. **11**, 103037 (2009).
  - [27] E. Hänggi, *Device-independent quantum key distribution*, PhD Thesis, Diss. ETH No. 19226, arXiv:1012.3878, (2010).

- [28] L. Masanes, S. Pironio, and A. Acín, *Secure device-independent quantum key distribution with causally independent measurement devices*, Nature. Commun. **2**, 238 (2011).
- [29] B. Toner, *Monogamy of nonlocal quantum correlations*, Proc. R. Soc. **465**, 59 (2009); A. Acín, N. Gisin, and L. Masanes, *From Bell's Theorem to Secure Quantum Key Distribution*, Phys. Rev. Lett. **97**, 120405 (2006).
- [30] N. Gisin *et al.*, *Trojan-horse attacks on quantum-key-distribution systems*, Phys. Rev. A. **73**, 022320 (2006).
- [31] M. Seevinck and J. Uffink, *Local commutativity versus Bell inequality violation for entangled states and versus non-violation for separable states*, Phys. Rev. A. **76**, 042105 (2007).
- [32] R. Renner, *Security of Quantum Key Distribution*. Diss. ETH No. 16242, arXiv:quant-ph/0512258, (2005).
- [33] A. E. Lita, A. J. Miller and S. W. Nam *Counting near-infrared single-photons with 95% efficiency*. Opt. Express **16**, 3032 (2008).
- [34] D. Stucki *et al.*, *Long-term performance of the Swiss Quantum quantum key distribution network in a field environment*, New J. Phys. **13**, 123001 (2011); M. Sasaki *et al.*, *Field test of quantum key distribution in the Tokyo QKD Network*, Opt. Express **19**, 10387 (2011).
- [35] G. Weihs *et al.*, *Violation of Bell's inequality under strict Einstein locality conditions*, Phys. Rev. Lett. **81**, 5039 (1998); W. Tittel *et al.*, *Long-distance Bell-type tests using energy-time entangled photons*, Phys. Rev. A. **59**, 4150 (1999); E. Pomarico *et al.*, *Various quantum non-locality tests with a commercial two-photon entanglement source*, Phys. Rev. A. **83**, 052104 (2011); M. S. Palsson, *et al.*, *Experimentally Violating Bell Inequalities Without Complete Reference Frames*, arXiv:1203.6692, (2012).
- [36] P. W. Shor and J. Preskill, *Simple Proof of Security of the BB84 Quantum Key Distribution Protocol*, Phys. Rev. Lett. **85**, 441 (2000).
- [37] M. N. Wegman and J. L. Carter, *New Hash Functions and Their Use in Authentication and Set Equality*, J. Comput. Syst. Sci. **22**, 265 (1981).
- [38] D. R. Stinson, *Universal Hashing and Authentication Codes*, Des. Codes Cryptogr. **4**, 369 (1994).
- [39] C. Portmann, *Key recycling in authentication*, arXiv:1202.1229, (2012).
- [40] J. Barrett, R. Colbeck and A. Kent, *Prisoners of their own device: Trojan attacks on device-independent quantum cryptography*, arXiv:1201.4407, (2012).
- [41] M. Pawłowski and N. Brunner, *Semi-device-independent security of one-way quantum key distribution*, Phys. Rev. A. **84**, 010302, (2011).
- [42] C. Jordan, *Essai sur la géométrie à n dimensions*, Bulletin de la S.M.F. **3**, 103 (1875).
- [43] L. Masanes, *Asymptotic violation of Bell inequalities and distillability*, Phys. Rev. Lett. **97**, 050503 (2006).
- [44] W. Hoeffding. *Probability Inequalities for Sums of Bounded Random Variables*, J. Amer. Statistical Assoc. **58**, 13 (1963).
- [45] R. J. Serfling. *Probability Inequalities for the Sum in Sampling without Replacement*, Ann. Stat. **2**, 39 (1974).
- [46] Sandu Popescu and Daniel Rohrlich. *Which states violate Bell's inequality maximally?*, Phys. Lett. A. **169**, 411, (1992); S. L. Braunstein, A. Mann, and M. Revzen, *Maximal violation of inequalities for mixed states*, Phys. Rev. Lett. **68**, 3259 (1992); V. Scarani and N. Gisin, *Spectral decomposition of Bell's operators for qubits*. J. Phys. A: Math. Gen. **34**, 6043 (2001).
- [47] A. De *et al.*, *Trevisan's extractor in the presence of quantum side information*, arXiv:0912.5514, (2009).
- [48] M. Tomamichel *et al.*, *Leftover Hashing Against Quantum Side Information*, IEEE Trans. Inf. Theory, **57**, 2703 (2011).